Substrate Discrimination by Formamidopyrimidine-DNA Glycosylase: Distinguishing Interactions within the Active Site†

Rebecca A. Perlow-Poehnelt,‡,§ Dmitry O. Zharkov,|| Arthur P. Grollman,⊥ and Suse Broyde*,‡

Department of Biology, New York University, 100 Washington Square East, Room 1009M, New York, New York 10003, SB RAS Institute of Chemical Biology and Fundamental Medicine, Novosibirsk 630090, Russia, and Department of Pharmaceutical Chemistry, State University of New York, Stony Brook, New York 11794

Received June 16, 2004; Revised Manuscript Received September 19, 2004

ABSTRACT: Reactive oxygen species are byproducts of normal aerobic respiration and ionizing radiation, and they readily react with DNA to form a number of base lesions, including the mutagenic 8-oxo-7,8-dihydroguanine (8-oxoG), 2,6-diamino-4-hydroxy-5-formamidopyrimidine (FapyG), 4,6-diamino-5-formamidopyrimidine (FapyA), and 8-oxo-7,8-dihydroadenine (8-oxoA). Such oxidative lesions are removed by the base excision repair pathway, which is initiated by DNA glycosylases such as the formamidopyrimidine-DNA glycosylase (Fpg) in *Escherichia coli*. The 8-oxoG, FapyG, and FapyA lesions are bound and excised by Fpg, while structurally similar 8-oxoA is excised by Fpg very poorly. We carried out molecular modeling and molecular dynamics simulations to interpret substrate discrimination within the active site of *E. coli* Fpg. Lys-217 and Met-73 were identified as residues playing important roles in the recognition of the oxidized imidazole ring in the substrate bases, and the Watson−Crick edge of the damaged base plays a role in optimally positioning the base within the active site. The recognition and excision of FapyA likely result from the opened imidazole ring, while 8-oxoA's lack of flexibility and closed imidazole ring may contribute to Fpg's inability to excise this base. Different interactions between each base and the enzyme specificity pocket account for differential treatment of the various lesions by this enzyme, and thus elucidate the structure−function relationship involved in an initial step of base excision repair.

Reactive oxygen species (ROS) are byproducts of normal aerobic respiration and ionizing radiation (*1*), and they are the major source of endogenous DNA damage in aerobic organisms (*2*). ROS readily react with DNA to form a number of covalently modified DNA bases. Some of the most abundant products of the reaction of ROS with DNA include 8-oxo-7,8-dihydroguanine (8-oxoG)[1] (*3*), 2,6-diamino-4-hydroxy-5-formamidopyrimidine (FapyG), 4,6-diamino-5-formamidopyrimidine (FapyA) (*4*−*6*), and 8-oxo-7,8-dihydroadenine (8-oxoA) (*6*, *7*) (see Figure 1 and the list of abbreviations for the respective 2′-deoxynucleosides of these damaged bases). 8-oxodG, an important mutagenic lesion produced by oxidative damage of DNA, primarily pairs with dC or dA during DNA replication (*8*). An 8-oxoG·A mispair

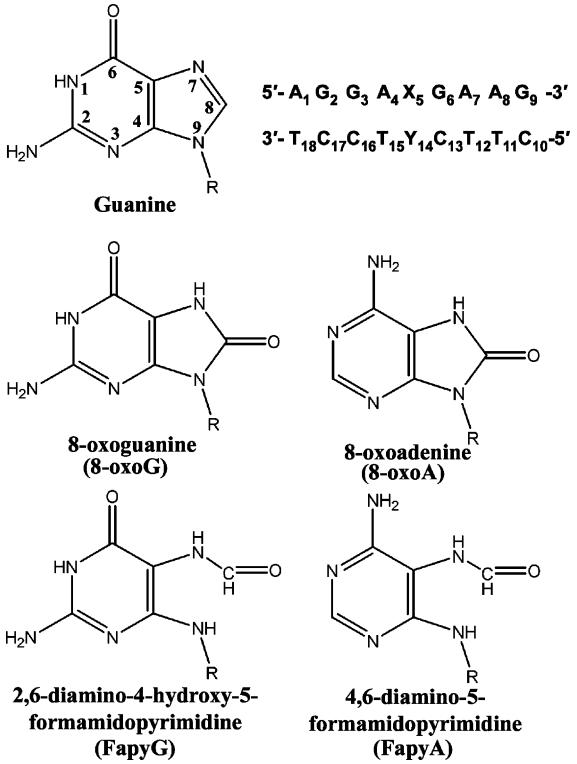

FIGURE 1: Chemical structure of the substrate bases and DNA sequence utilized in the Fpg molecular dynamics studies. Each substrate base was positioned at X_5, while the normal partner to its undamaged counterpart was modeled at the Y_{14} position.

† This work has been supported by American Cancer Society Grant PF-01-108-01-CNE (R.A.P.-P.), the Wellcome Trust (U.K.) and the Russian Foundation for Basic Research (Grant 02-04-49605) (D.O.Z.), NIH Grant CA-17395 (A.P.G.), and NIH Grant CA-75449 (S.B.).

* To whom correspondence should be addressed. E-mail: broyde@nyu.edu.

‡ New York University.

§ Current address: Merck Research Laboratories, Merck & Co., Inc., P.O. Box 4, Sumneytown Pike, West Point, PA 19486.

|| SB RAS Institute of Chemical Biology and Fundamental Medicine.

⊥ State University of New York.

[1] Abbreviations: 8-oxoG, 8-oxo-7,8-dihydroguanine; FapyG, 2,6-diamino-4-hydroxy-5-formamidopyrimidine; FapyA, 4,6-diamino-5-formamidopyrimidine; 8-oxoA, 8-oxo-7,8-dihydroadenine; 8-oxodG, FapydG, FapydA, and 8-oxodA, respectively 2′-deoxynucleosides of these lesions; Fapy, formamidopyrimidine; Fpg, formamidopyrimidine-DNA glycosylase from *Escherichia coli*; Bst-Fpg, Fpg from *Bacillus stearothermophilus*.

10.1021/bi048747f CCC: $27.50
Published on Web 11/25/2004

results in a G → T transversion when the damaged base is on the template strand, while an A → C transversion can result when a modified base is introduced into DNA through pairing of its nucleotide triphosphate with a template dA during replication (*9*−*11*). 8-oxodA lesions are also mutagenic in cells, and the most common mispair is with dG which produces A → C transversion mutations (*12*). Both FapyG and FapyA can also facilitate mismatches that may lead to mutations (*13*, *14*).

Cells employ a number of pathways by which DNA damage is removed to avert mutagenic consequences. Small base modifications, such as those resulting from oxidation, alkylation, or deamination, are often repaired by the base excision repair (BER) pathway (*15*). BER is initiated by DNA glycosylases, which recognize damaged and, in some cases, mismatched bases and remove them from DNA by cleaving the *N*-glycosidic bond between the base and deoxyribose. The free base is then released, leaving an apurinic/apyrimidinic (AP) site in the duplex that is further processed by the BER machinery to reconstitute the original undamaged DNA sequence (*4*, *16*).

In *Escherichia coli*, repair of a number of oxidative lesions, including 8-oxodG (*17*), FapyG (*5*), and FapyA (*5*, *18*), is initiated by the formamidopyrimidine-DNA glycosylase (Fpg). Fpg specifically distinguishes its oxidatively damaged substrate bases from the vast quantity of undamaged dG and dA, as well as other damaged bases, within the genome. Once a damaged substrate base is selected by Fpg, it not only cleaves the damaged base moiety from the sugar−phosphate backbone but also engages in an associated lyase activity, during which it cleaves the damaged strand to remove the abasic site (*19*). The fact that these lesions all cause helix destabilization may be a factor in their repair susceptibility (*13*, *20*−*23*).

The 8-oxoG (*17*, *24*), FapyG (*25*), and FapyA (*20*) bases are all bound with high affinity and excised by the Fpg protein, and Fpg shows a preference for the partner base opposite 8-oxoG: T > G > C ≫ A (*17*). The discrimination against A opposite 8-oxoG likely evolved to avoid excising the mismatched damaged base, leaving the A to fix the G → T transversion mutation when the resulting gap is filled in during subsequent steps of BER (*26*). Interestingly, Fpg does not remove the structurally similar 8-oxodA lesion from duplex DNA (*17*). Unmodified guanine is also not cleaved by Fpg to a significant extent (*17*).

The substrate specificity of the 30 kDa Fpg protein was investigated using oligodeoxynucleotides containing bases that were systematically modified for delineation of the importance of various functional groups of substrate bases to binding and excision by the enzyme. Modified bases each fall into one of three categories: (1) those that are neither bound relatively well nor cleaved by Fpg, including unmodified bases and 8-methoxyguanine (*17*; D. O. Zharkov, unpublished results); (2) those that are bound relatively well by the enzyme but are not efficiently removed from the damaged strand, namely, 8-oxonebularine, 6-*O*-methyl-8-oxoguanine, and, possibly, 8-oxoA (*17*; D. O. Zharkov, unpublished results); and (3) those that are bound and excised by Fpg, namely, 8-oxoG, 8-oxoinosine (*17*), FapyG (*25*), and FapyA (*20*). While the 2-amino group does not appear to play a role in binding or cleavage efficiency, the keto group at position 8, with its accompanying pyrrolic nitrogen at position 7, is required for binding of Fpg to the substrate base, presumably used to distinguish damaged from undamaged bases. In addition, a keto group at position 6 of the substrate base, with its accompanying pyrrolic nitrogen at position 1, appears to signal for cleavage of the base by Fpg. One exception to this rule is FapyA, which contains an amino group at C6 rather than a keto group, as well as a pyridinic nitrogen at position 1, but is efficiently bound and cleaved by Fpg (*27*). It is possible that the additional conformational flexibility of FapyA, conferred by its ring-opened structure, allows the base moiety to adopt a conformation that satisfies the enzyme's substrate requirements; the lack of this flexibility may explain why 8-oxodA is not cleaved by Fpg.

Productive substrate binding by Fpg involves several rapid conformational changes that facilitate catalysis (*24*). The Fpg protein catalyzes substrate base removal and strand scission through a reaction pathway that entails formation of a Schiff base between the N-terminal proline (Pro-1) of the enzyme and the damaged DNA strand (*19*). A stable covalent protein−DNA complex can be trapped by reduction of the Schiff base intermediate with $NaBH_4$ (*28*), and the crystal structure of this Fpg−DNA complex reduced Schiff base has been determined (*29*). Although a damaged base is not present in the crystal structure, the complex reveals the orientation of DNA binding and a possible position for the substrate base. The Fpg-bound DNA resides in a predominantly basic crevice of the protein and is kinked by approximately 66°, centered upon the lesion site. In addition, the minor groove of the duplex, which is primarily contacted by the protein, is widened significantly. This structure suggests that the substrate base is flipped out of the helix because of the extruded ring-opened sugar moiety and an adjacent putative specificity pocket (*29*). Such a flipped-out conformation of the damaged base is not unexpected, as similar phenomenona occur in other DNA glycosylase enzymes, including human 8-oxoguanine-DNA glycosylase Ogg1 (*30*), 3-methyladenine-DNA glycosylase (*31*), uracil-DNA glycosylase (*32*), *E. coli* AlkA protein (*33*), and a catalytically inactive *Bacillus stearothermophilus* Fpg (Bst-Fpg) (*34*). The gap that results from substrate base extrusion is filled by hydrophobic protein residues Met-73 and Phe-110, as well as Arg-108, which also participates in hydrogen bonds with the widowed partner of the displaced substrate base. The crystal structure suggests that the damaged base may reside in a binding pocket whose boundaries are defined by Pro-1, Glu-2, Glu-5, Ile-59, Tyr-170, Thr-214, Thr-215, and Leu-216 (*29*, *35*), and molecular modeling, molecular dynamics, and mutational studies implicate Lys-217 and Met-73 in the recognition of the O^8 and N7 moieties of the 8-oxodG substrate (*35*). Despite extensive functional and structural characterization, the mechanism by which Fpg discriminates between structurally similar DNA bases remains to be elucidated; to date, there are no crystal structures of the catalytically active complexes simulated in this study.

We carried out molecular modeling and molecular dynamics simulations, as well as detailed structural analyses, to elucidate the mechanism of substrate discrimination within the active site of *E. coli* Fpg. Specifically, we investigated the structural distinctions between Fpg recognition of guanine, 8-oxoA, 8-oxoG, FapyG, and FapyA. These lesions were selected because they represent a range of recognition and excision susceptibilities for which experimental data are

available. Different interactions between each base and the enzyme specificity pocket elucidate a mechanism of substrate discrimination for the Fpg repair enzyme.

MATERIALS AND METHODS

Simulations were carried out using the SANDER module of the AMBER 5.0 molecular modeling and molecular dynamics software package (*36*), the force field of Cornell et al. (*37*), and the parm99.dat parameter set (*38*). Nonbonded interactions were updated every time step, and a 9 Å cutoff was applied to Lennard-Jones interactions; periodic boundary conditions were applied, and MD simulations were carried out under constant pressure (1 atm) and temperature (300 K) unless indicated otherwise. Coulombic interactions were approximated using the particle mesh Ewald method (*39*); bonds involving hydrogen atoms were constrained using the SHAKE option with a tolerance of 0.0005 Å, and a 2 fs time step was used throughout the simulation. Coordinates of the crystal structure of the Fpg protein covalently linked to DNA [Protein Data Bank entry 1K82 (*40*)], including a homology-modeled eight-residue flexible loop (residues 217−224) that was not located in the crystal structure, were obtained from Gilboa et al. (*29*, *35*). The N-terminal proline residue, which acts as the nucleophile in Fpg's catalysis of substrate base glycosidic bond cleavage, was modeled as being neutral to mimic the scenario directly preceding the reaction.

Residues Not Included in the AMBER Force Field Were Parametrized. Partial charges for the neutral N-terminal Pro-1 (PRN), 8-oxodG, 8-oxodA, FapydG, and FapydA residues were calculated for use in parametrization of these residues. The conformation of the Pro-1 residue in the crystal structure was used in the parametrization of the PRN residue. For parametrization of 8-oxodG, three conformations that differed in their glycosidic torsion angle were used to obtain partial charges (*35*); χ torsions of 255°, 294°, and 317° were employed. The base moiety in the NMR structure of the formamidopyrimidine adduct of aflatoxin B1 (*41*) was used as a starting structure for FapydA and FapydG conformations in the partial charge calculation. The aflatoxin moiety was removed, and the base was remodeled using the Builder module of InsightII (Accelrys, Inc.) to either FapydA or FapydG to yield the structures used in the calculations. For all residues being parametrized, Hartree−Fok calculations with the 6-31G* basis set in GAUSSIAN 98 (*42*) were used to calculate the partial charge of each atom. The RESP module of AMBER was used to fit each charge to its respective atomic center (*43*), and charges were normalized to prevent charge imbalance in each residue (*44*). The final partial charges and atom type assignments for the PRN, 8-oxodG, 8-oxodA, FapydG, and FapydA residues are given in Tables S1−S5 of the Supporting Information. Bond and angle parameters not included in the original force field were assigned by analogy to chemically similar types in the AMBER database parm99.dat parameter set and are listed in Table S6 of the Supporting Information.

Starting Structures Were Prepared Using Molecular Modeling and Equilibration Was Carried Out Prior to Molecular Dynamics Production Runs. The 2.1 Å resolution crystal structure of Fpg covalently linked to DNA through the reduced Schiff base intermediate was used as the starting structure for models, and the coordinates were obtained from Gilboa et al. (*29*). The coordinates included a homology-modeled loop of residues 217−224 that was not located in the original crystal structure due to its high mobility. Molecular modeling was used to break the covalent bond between the Pro-1 residue and ring-opened sugar moiety comprising the reduced Schiff base, and the intact 2′-deoxyribose moiety was reconstructed. The modeled sequence corresponded to that in the crystal structure, 5′-AGGA**N***GAAG-3′, where **N*** denotes the substrate base. Four starting structures of 8-oxodG within the Fpg active site were prepared that differed only in the glycosidic torsion angles of the substrate base (*35*); three starting structures had *anti*−*high anti* glycosidic torsion angles (χ) of the damaged base, while one had a *syn* glycosidic torsion angle. The glycosidic torsion angle, χ, is defined as the O4′−C1′−N9−C4 (purine) and O4′−C1′−N1−C2 (pyrimidine) torsion angle. Starting structures were chosen to maximize stabilizing protein−DNA contacts (*35*). *Anti* 8-oxodG simulations 1−3 had starting χ torsion angles of 294°, 317°, and 255°, respectively. Simulations 1 and 3 converged to give very similar results that were consistent with experimental findings in mutagenicity studies (*35*), supporting the significance of these conformations. The *syn* 8-oxodG simulation had a starting χ torsion angle of 110°, and this simulation also gave results consistent with experimental findings (*35*). Thus, starting glycosidic torsions of 294° and 110° were chosen for *anti* and *syn* simulations, respectively, of the remaining substrates, namely, unmodified guanine, FapydG, FapydA, and 8-oxodA, to parallel the 8-oxodG simulations.

Once the DNA was remodeled, 750 steps of steepest descent (SD) and 250 steps of conjugate gradient (CG) potential energy minimization were carried out using the SANDER module of AMBER 5.0. All minimizations prior to the addition of the periodic box of solvating waters were carried out using the Hingerty distance-dependent dielectric function (*45*) implemented in our version of AMBER 5.0. Following minimization, 425 crystallographic water molecules were added to each system, and the systems were minimized for 600 steps of SD, while holding the solute (protein, DNA, and Zn^+) fixed using harmonic restraints of 0.1 kcal mol^{-1} Å^{-1}. A total of 20 neutralizing Na^+ counterions were then added to each system at positions of lowest electrostatic potential using the LEaP module of AMBER 5.0. The neutralized systems were minimized for 1000 steps of SD while holding the solute fixed with harmonic restraints of 10 kcal mol^{-1} Å^{-1}. The systems were reoriented using the Simulaid program (*46*) and then solvated in a box of TIP3P waters using a buffer of 10 Å. The solvating waters, crystallographic waters, and counterions were minimized for 5000 steps of SD while keeping the solute fixed with 10 kcal mol^{-1} Å^{-1} restraints, with a constant dielectric function of 1. MD was carried out to equilibrate the system: 30 ps at 10 K while holding the solute fixed with 10 kcal mol^{-1} Å^{-1} restraints; the system was heated for 25 ps from 10 to 300 K while holding the solute fixed with 10 kcal mol^{-1} Å^{-1} restraints, for 45 ps at 300 K while holding the solute fixed with 10 kcal mol^{-1} Å^{-1} restraints, for 20 ps at 300 K while holding the solute fixed with 1 kcal mol^{-1} Å^{-1} restraints, and for 30 ps at 300 K while holding the solute fixed with 0.1 kcal mol^{-1} Å^{-1} restraints. Following equilibration for

150 ps, unrestrained production MD was carried out at 300 K for 2 ns.

Coordinates, velocities, and energies for each simulated system were collected every picosecond. Each trajectory was analyzed to pinpoint interactions between the protein and DNA that could help elucidate the mechanism of substrate discrimination. The CARNAL and PTRAJ modules of AMBER were used for the trajectory processing and analyses of the atomic fluctuations, critical torsion angles and distances, interactions between the solute and solvent, and hydrogen bond occupancy and quality involving the glycosylase and substrates. The hydrogen bonding cutoff for the heavy atom−heavy atom distance was 3.4 Å, and that for the hydrogen bonding angle was 140°.

RESULTS AND DISCUSSION

Four simulations of 8-oxodG, in addition to two each of unmodified guanine, FapydA, FapydG, and 8-oxodA, were carried out for a total of 12 simulations. Each base was simulated in both the *anti* and *syn* conformations within the active site region of *E. coli* Fpg, and stereoviews of the active site of each complex after unrestrained MD for 2 ns are shown in Figure 2; *anti* 8-oxodG simulation 3 is not shown since it converged to a conformation very similar to that of *anti* 8-oxodG simulation 1.

The Substrate Bases Are Stably Accommodated within the Active Site Region, Flipped Out from the Duplex DNA. Each system had a relatively low rmsd from the starting structure, demonstrating the stability of the simulations. Figure S1 shows the rmsd of the protein and DNA in each system from the starting structure over the 2 ns simulations; the average rmsd, with the standard deviation, is given within each of these plots. The insets show the two-dimensional rmsd analysis of the substrate base. The substrate base is accommodated stably within the active site region of each system, including simulations in which the substrate was *anti* and *syn*. Although the damaged bases can be accommodated in the *syn* conformation within the Fpg active site, the 8-oxoG and FapyG substrate bases that started in the *syn* orientation rotated significantly about their glycosidic bonds toward the *anti* domain. Figure S2 shows the χ torsion angles of the substrate bases in each simulation.

Fpg Recognition of Oxidatively Damaged Substrate Bases within the Active Site Is Carried Out by Lys-217 and/or Met-73. The Fpg protein is involved in a number of contacts with its DNA substrate, including numerous nonspecific interactions with the phosphodiester backbone, as revealed in the crystal structure of *E. coli* Fpg covalently linked to DNA through a reduced Schiff base intermediate (*29*). Modeling substrate bases into the active site of Fpg and subsequent MD simulations allow the observation of specific interactions between the Fpg protein and the base moieties themselves. Including substrate bases that are structurally distinct, namely, 8-oxoG, FapyG, and FapyA, as well as bases that are not cleaved by the enzyme, namely, unmodified guanine and 8-oxoA, allows the comparison of protein−DNA interactions in each system that elucidate the activity of Fpg at each base.

We previously carried out an analysis of 8-oxodG within the active site of Fpg to guide site-specific mutagenesis experiments, and we predicted correctly that Lys-217 is involved in substrate discrimination (*35*). In *anti* 8-oxodG simulation 1, the O^8 group of the damaged base is recognized through a stable hydrogen bond with Nζ of Lys-217 of the Fpg protein, as illustrated in Figure 2, with a distance of 3.0 ± 0.7 Å. The length of this hydrogen bonding interaction briefly increases significantly near the end, but then again regains hydrogen bonding distance before the end of the simulation; this temporary disruption does not occur in *anti* 8-oxodG simulation 3, which converges to a conformation very similar to that of simulation 1. In addition, when 8-oxodG is modeled into the *syn* conformation within the active site, the substrate base not only rotates about its glycosidic torsion angle to adjust its position within the active site but also participates in a hydrogen bond with Nζ of Lys-217 for 85% of the simulation, with a distance of 3.0 ± 0.4 Å; this hydrogen bond is not within reach at the beginning of the simulation, but only forms upon adjustment of the 8-oxodG base within the active site. The distances and angles critical to damage recognition for each system are shown in Figure 3. The O^8 group is a hydrogen bond acceptor that distinguishes 8-oxodG from unmodified guanine, which contains an unsubstituted aromatic carbon at the 8 position. Indeed, mutation of Lys-217 to Thr decreased the activity of Fpg on 8-oxodG significantly, but did not affect the activity on a structurally different Fpg substrate, namely, 5,6-dihydrouracil (*35*). Although this hydrogen bond between 8-oxodG O^8 and Nζ of Lys-217 is absent in 8-oxodG simulation 2, a hydrogen bond forms between the main chain O of Met-73 and N7 of the 8-oxodG moiety. Since this hydrogen bond involves the main chain oxygen of Met-73, its influence on substrate discrimination cannot be tested by site-directed mutagenesis. It is possible that if the O of Met-73 were replaced with a group that was incapable of acting as a hydrogen bond acceptor, the 8-oxoG base moiety would rearrange within the active site to interact with Lys-217 through its O^8 group. Figure 4 illustrates the occupancy of the hydrogen bonds involving the substrate base moiety and the active site region for each simulation.

Interestingly, FapyG and FapyA, two structurally distinct formamidopyrimidine lesions, purine derivatives with opened imidazole rings, can also be recognized and cleaved by Fpg. Although the O^8 groups of these two bases do not participate in stable hydrogen bonds with the enzyme, the oxidative damage-specific anilide N7 and amino N9 of the open imidazole ring are involved in stable interactions with Met-73 during the simulations when the bases adopt an *anti* conformation. FapyG participates in two water-mediated hydrogen bonds involving crystallographic water 379: one between N9 of FapyG and the O atom of Met-73 and another between N7 of FapyG and the O atom of Met-73. The N9 and N7 atoms of FapyG hydrogen bond with crystal water 379 for 100% of the simulation, while the water molecule hydrogen bonds with the O atom of Met-73 for 99% of the simulation. Water-mediated hydrogen bonds are often involved in specific protein−DNA interactions (*47*). Similarly, N7 and N9 of FapyA interact with Met-73, participating in hydrogen bonds with the main chain oxygen of Met-73 for 92 and 63% of the simulation, respectively. The distances and angles of these interactions are given in Figure 3. In contrast, when either FapydG or FapydA is modeled in the *syn* conformation, the damaged bases hydrogen bond with neither Met-73 nor Lys-217 during the simulations; it appears

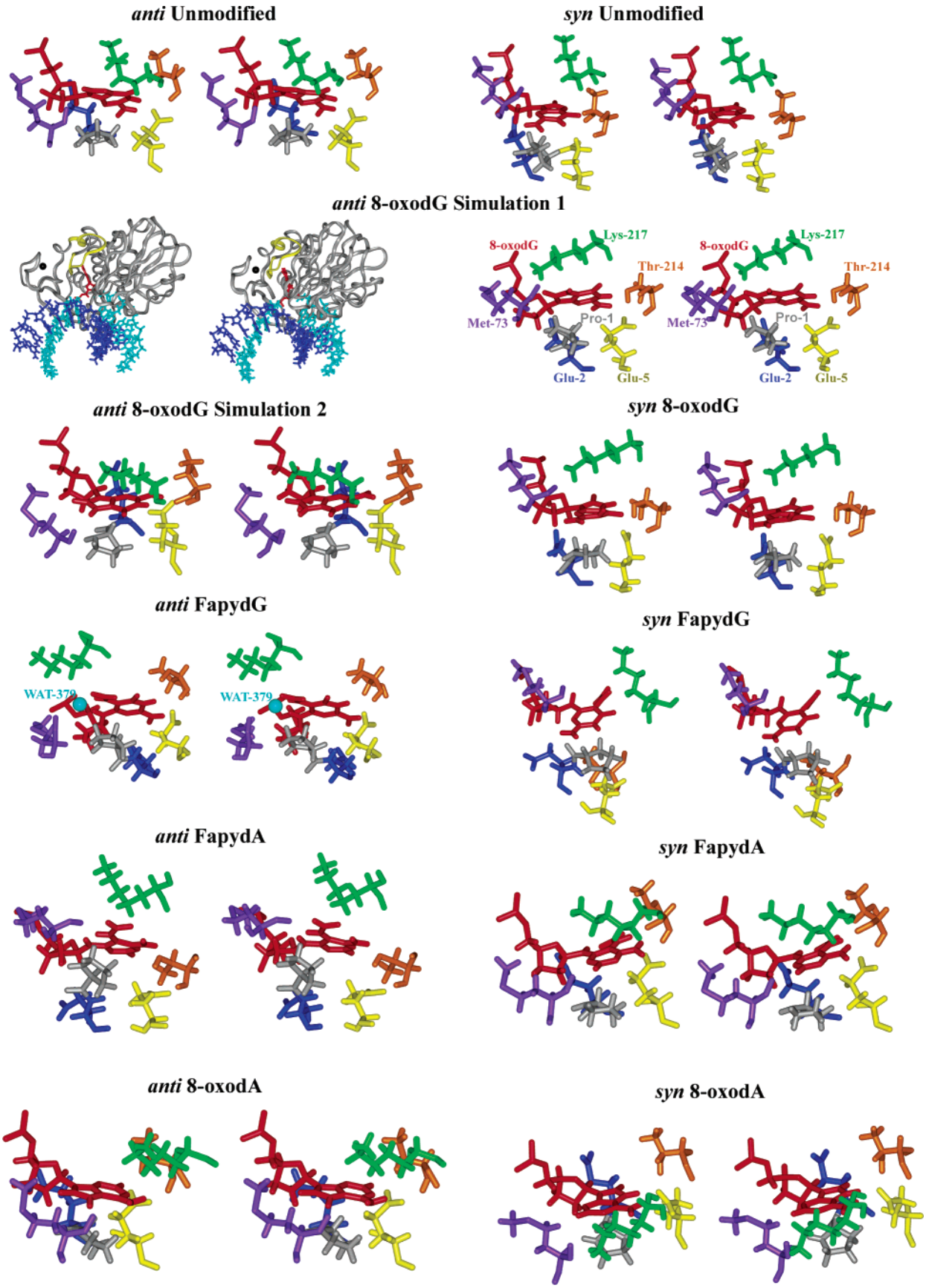

FIGURE 2: Stereoviews of the damaged base and critical active site residues in each simulation. A stereoview of the entire protein−DNA complex for *anti* 8-oxodG simulation 1 shows the relative position of the flipped-out base, with 8-oxodG in red. The color code for the stereoview of the entire complex follows: gray for the ribbon model of Fpg, yellow for the flexible loop region (residues 215−228), dark blue for the damaged DNA strand, cyan for the partner undamaged DNA strand, red for 8-oxodG, and black for Zn^{2+}. The amino acid color code for the active site stereoviews is shown for *anti* 8-oxodG simulation 1, and Wat-379 is shown as cyan in the *anti* FapyG structure. Stereoviews were made for viewing with a stereoviewer.

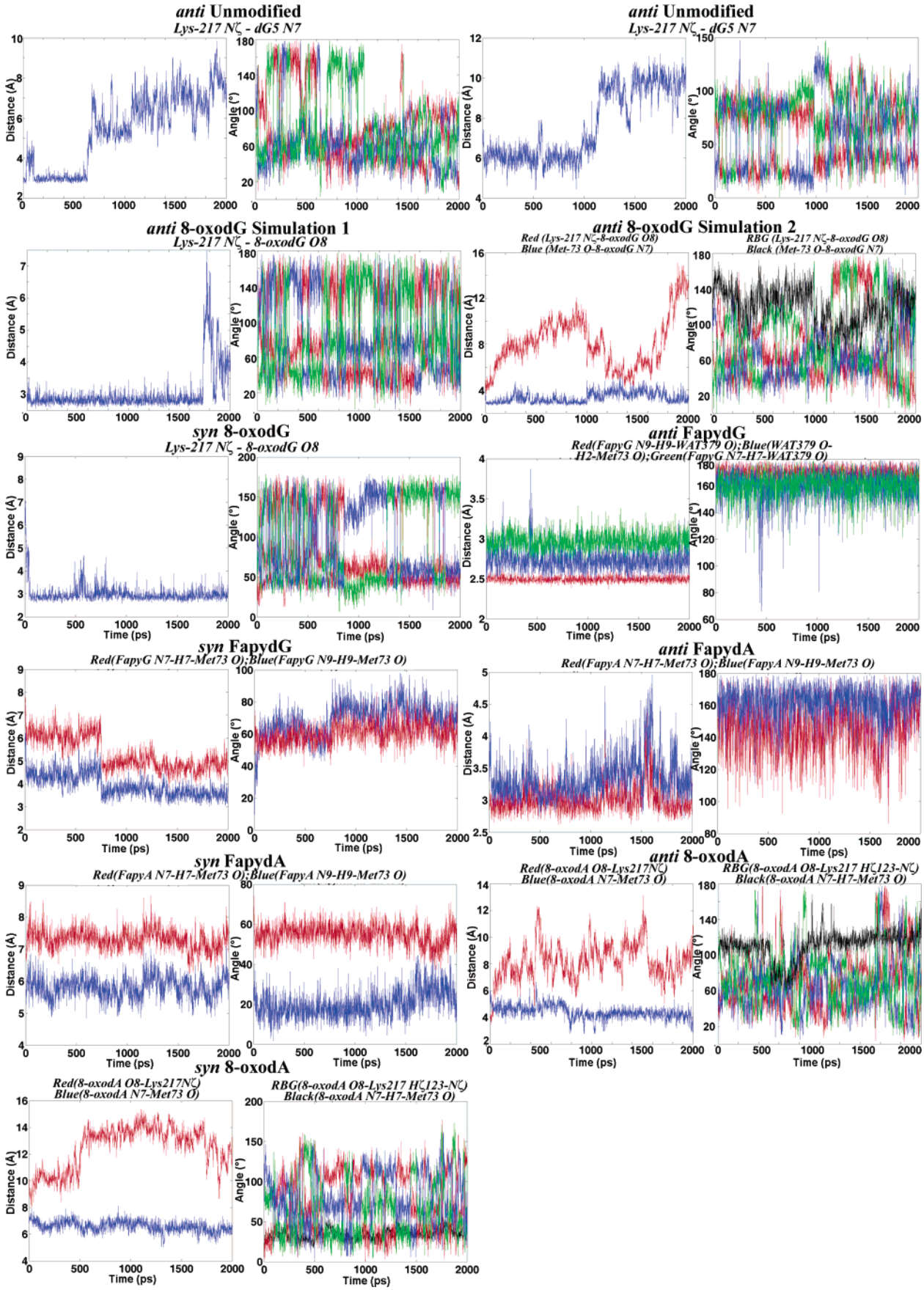

FIGURE 3: Critical hydrogen bonding distances and angles involved in recognition of the O^8 and/or N7/N9 groups in each simulation. Three hydrogen bonding angle values are measured for each hydrogen bonding interaction with Lys-217 because the Hζ atoms freely rotate during these simulations. RBG indicates red, blue, and green.

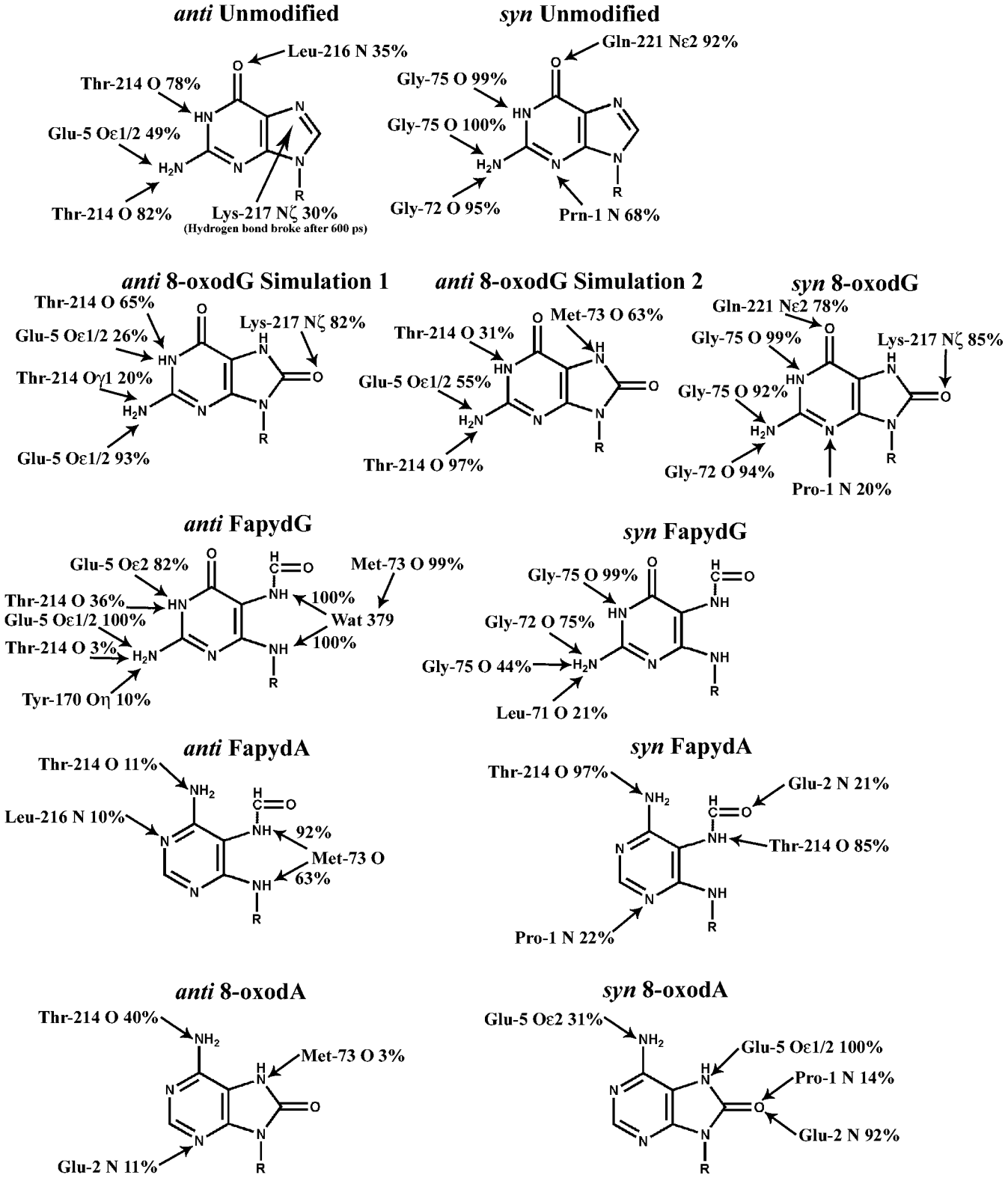

FIGURE 4: Schematic diagram of the hydrogen bonding interactions involving the substrate base moiety during the simulations. The percent occupancy is given for each interaction, defined as a heavy atom−heavy atom distance of ≤3.4 Å and a heavy atom−hydrogen−heavy atom angle of ≤140°.

that while 8-oxodG may be recognized by Fpg in either the *anti* or *syn* conformation, FapydG and FapydA must adopt the *anti* conformation to be specifically recognized within the active site of the enzyme. The 7-methylformamidopyrimidine base (Me-Fapy) contains a methyl group at position N7 of the imidazole ring, but is also a very good substrate for Fpg (*48*, *49*). It is likely that in the case of Me-Fapy, the interaction of the N9 amino group with Met-73 is sufficient to facilitate recognition. Alternatively, the extra bulk of the methyl group on N7 of Me-Fapy could cause the rearrangement of the damaged base so that its O^8 group is recognized by Lys-217.

Proper positioning of the substrate base within the active site facilitates the alignment of reactive groups for catalysis.

In Fpg, Pro-1 carries out a nucleophilic attack on C1′ of the substrate base (*28*), and thus, these two groups must be adjacent for base excision to occur. Figure S3 shows the distance from the N atom of Pro-1 to C1′ of the damaged nucleotide for each simulation. For each of the three excised substrates, 8-oxodG, FapydG, and FapydA, this distance is longer in the simulations in which the damaged nucleoside adopts the *syn* conformation. In the *anti* conformations of the substrate nucleosides, the Pro-1−C1′ distance at least samples within a range reasonable for the subsequent reaction (ca. <3.5 Å). A shorter Pro-1 N−C1′ distance is one indicator of proper accommodation within the active site pocket.

Unmodified Guanine and 8-oxoA which Are Not Cleaved by Fpg Do Not Participate in Stable Specific Interactions with Lys-217 or Met-73. Although the oxidized Hoogsteen edge of the 8-oxoG, FapyG, and FapyA bases is specifically recognized by Lys-217 and/or Met-73, this same profile in 8-oxoA is not recognized within the active site of Fpg. Significant hydrogen bonds between the protein and damaged base do not form in either of the 8-oxodA simulations, whether the base is *anti* or *syn*. In the *anti* 8-oxodA simulation, the N7 group of the damaged base is hydrogen bonded to the O atom of Met-73 only transiently, for only 3% of the simulation, likely well below the threshold required to signal for substrate base cleavage.

Likewise, unmodified guanine does not participate in stable hydrogen bonds with Lys-217 or Met-73 of Fpg in either the *anti* or *syn* conformation. The pyrrolic N7 atom of an oxidized purine, able to act as a hydrogen bond donor, is distinct from the pyridinic N7 atom in unmodified guanine, which can act as a hydrogen bond acceptor. Since N7 is a hydrogen bond acceptor in guanine, like O^8 in 8-oxoG, it is possible that this group could interact with Lys-217 also, making the base appear to be damaged. In the *syn* orientation, N7 of unmodified guanine does not participate in any hydrogen bonds with the active site region. However, in the *anti* unmodified simulation, the substrate base partakes in a hydrogen bond with Lys-217 for the first 600 ps of the simulation with a hydrogen bonding distance of 3.1 ± 0.4 Å; however, this hydrogen bond breaks, and the two groups remain well outside of hydrogen bonding distance for the remainder of the simulation, as shown in Figure 3. The distance between Nζ of Lys-217 and N7 of guanine in the *anti* unmodified simulation is 6.4 ± 1.2 Å for the final 1400 ps of the simulation, while in the *syn* conformation, this distance never reaches less than ∼4.5 Å. N7 of unmodified guanine does not attain an orientation that allows a stable interaction with Lys-217 due to its different placement on the base compared to O^8 of 8-oxoG.

The simulations of Fpg complexed to bases that it excises from duplex DNA, including 8-oxoG, FapyG, and FapyA, as well as bases that it does not cleave from duplex DNA, namely, unmodified guanine and 8-oxoA, point to Lys-217 and Met-73 as being critical in substrate discrimination within the active site pocket. It is likely that a stable interaction of 8-oxodG O^8 and Lys-217 Nζ or 8-oxoG N7 and Met-73 O is required to trigger cleavage of the glycosidic bond and subsequent steps of Fpg-mediated repair of the damaged DNA. In addition, interaction with N7 and N9 in the imidazole ring-opened formamidopyrimidine residues, FapyG and FapyA, appears to be necessary for substrate cleavage, whether these interactions are direct or water-mediated. Similarly, a crystal structure of 8-oxodG bound by Fpg's functional analogue in humans, hOgg1, showed that specific recognition of the substrate base hinged on a single hydrogen bond (*30*). The damaged base adopts the *anti* conformation within the active site of hOgg1, and the critical discriminating hydrogen bond is formed between the damage-specific pyrrolic N7 atom on the Hoogsteen edge of the lesion and a protein main chain carbonyl group.

Interactions between Fpg and the Watson−Crick Edge of 8-oxoG May Help Position the Oxidized Hoogsteen Edge for Interaction with the Damage Recognition Machinery. The Watson−Crick edge of the substrate base appears to play an important role in signaling for cleavage by Fpg. While the O^8 group is critical for binding of Fpg to the damaged base (*48*), the presence of the carbonyl group at C6 and/or its associated pyrrolic N1 signals for the enzyme to excise the damaged base from the duplex (*17*). It is also possible that the Watson−Crick edge of the substrate base plays a more indirect role in signaling for cleavage, by aiding in the positioning of the base within the active site region, facilitating the Hoogsteen edge's interaction with Lys-217 and/or Met-73 that directly signals for cleavage. Figure 4 shows a schematic depiction of the hydrogen bonds between the substrate base and the Fpg active site region during the simulations, including their occupancies, defined as the percentage of the simulation in which the interaction has a hydrogen bonding distance of ≤3.4 Å (heavy atom−heavy atom distance) and an angle of ≥140° (heavy atom−hydrogen−heavy atom angle).

In the *syn* unmodified simulation, O^6 of guanine is involved in a hydrogen bond with Gln-221 for 92% of the simulation, but N1 does not participate in any hydrogen bonds with the protein. In addition, the *syn* unmodified guanine's N^2 atom hydrogen bonds to Gly-72 and Gly-75 for >95% of the simulations. The Gln-221, Gly-75, and Gly-72 residues also interact with the damaged base in the *syn* 8-oxodG simulation, in a pattern similar to that seen in the *syn* unmodified system. In contrast, the *anti* unmodified guanine's N1 participates in a hydrogen bond with the O atom of Thr-214 for 78% of the 2 ns simulation, while O^6 of the substrate base hydrogen bonds to the N atom of Leu-216 for 35% of the simulation. The NH_2 group of the *anti* unmodified guanine hydrogen bonds to Glu-5 and Thr-214 for 49 and 82% of the simulation, respectively. The damaged base in the *anti* 8-oxodG simulation exhibits a pattern of hydrogen bonding similar to that of the *anti* unmodified guanine, although it lacks an interaction between its O^6 group and the protein. As shown in Figure 4, Glu-5 hydrogen bonds to N1 of the *anti* 8-oxodG base for 26% of the simulation while the O atom of Thr-214 hydrogen bonds to N1 of the damaged base for 65% of the simulation. In addition, the NH_2 moiety of the damaged base in the *anti* 8-oxodG simulation partakes in hydrogen bonds with Glu-5 for 93% of the simulation and with Oγ1 of Thr-214 for 10% of the simulation.

The NH_2 group of 8-oxodG has been deemed unnecessary for processing by Fpg (*17*) due to the unhindered excision of 8-oxoinosine, which lacks an amino group at the 2 position, from duplex DNA. However, an altered hydrogen bonding profile for this group reflects the differing position of the substrate base and/or rearrangement of the active site

region in the *anti* 8-oxodG simulation 1 compared to the unmodified systems. The hydrogen bonding profile on the Watson−Crick edge of the substrate base in both *anti* 8-oxodG simulation 2 and *anti* FapyG parallels that of *anti* 8-oxodG simulation 1, but a similar hydrogen bonding profile is not reflected in the *syn* FapyG simulation, as shown in Figure 4. In addition, the substrate base in the *syn* FapyG simulation interacts with neither Lys-217 nor Met-73, suggesting that the FapyG base may be more readily recognized when it adopts the *anti* conformation.

Although the Watson−Crick edge hydrogen bonding profile in the *anti* unmodified simulation is similar to that in *anti* 8-oxodG simulation 1, the small differences, such as Leu-216 hydrogen bonding to O^6 and the failure of Glu-5 to hydrogen bond to N1 in the unmodified system, help explain the differences in Fpg activity on these two bases. While the Watson−Crick edge interactions help position the substrate base for interaction with Lys-217 and Met-73, the orientation of N7 in the unmodified base may not be ideal for this interaction, resulting in this interaction being severed after a fraction of the simulation. Unmodified guanine is bound very poorly by Fpg, indicating that the base is less frequently within the active site of the enzyme. However, due to the abundance of guanine in the genome compared to substrate lesions, at least a portion of the discrimination between unmodified guanine and damaged bases is carried out within the active site pocket of the enzyme; the discriminating protein−base interactions pinpointed in these simulations could explain Fpg's relative inability to excise unmodified guanine (*17*).

In contrast, the Watson−Crick edges of 8-oxoA and FapyA differ from that of their guanine counterparts. While this difference adds to the rationalization of the failure of Fpg to excise 8-oxoA, it adds to the mystery of how FapyA is removed by the enzyme. The substrate bases in the *anti* and *syn* 8-oxodA simulations do not partake in interactions with either of the damage site recognition residues, Met-73 or Lys-217, and this shortcoming, which likely causes the inability of Fpg to excise 8-oxoA, may result from the lack of appropriate interactions of the Watson−Crick edge of the damaged base, namely, those involving the 1 and 2 positions, with Thr-214 and Glu-5. Instead, hydrogen bonds exist between N^6 of *anti* 8-oxodA and the O atom of Thr-214 for 40% of the simulation, as well as between N^6 of *syn* 8-oxodA and Glu-5 for 31% of the simulation. Although the 8-oxodA base contains the correct recognition moieties on its Hoogsteen edge, the failure of its Watson−Crick edge to forge the requisite interactions with the active site region likely results in the misalignment of the O^8 and/or N7 groups with Met-73 and Lys-217. Interestingly, FapyA possesses a Watson−Crick edge identical to that of 8-oxoA, yet FapydA is both bound with high affinity and excised by Fpg. Although N^6 of the substrate base in the *anti* FapydA simulation is hydrogen bonded to Thr-214 for 11% of the simulation and the Watson−Crick edge of the substrate base does not interact with Glu-5, N7 and N9 are involved in hydrogen bonds with the O atom of Met-73 for 92 and 63% of the simulation, respectively. It is possible that the flexibility and different hydrogen bonding potential of the imidazole ring-opened FapyA base allow it to forge interactions with Fpg that signal for base cleavage and appropriately align the damaged base within the active site, while these interactions are not accessible to the conformationally restricted 8-oxoA base. The FapyA base also partakes in a hydrogen bond with Thr-214 through its N^6 group in the *syn* conformation, but the substrate base in this simulation does not interact with Met-73 or Lys-217; it is quite possible that, like FapyG, FapyA is also excised solely when it adopts the *anti* orientation within the active site of Fpg.

The Position of the Substrate Base within the Active Site of Fpg Explains the Loss of Glycosylase Activity Observed Experimentally for a Number of Mutant Enzymes. Site-directed mutagenesis studies have pinpointed a number of residues that are critical to the glycosylase activity of Fpg. The N-terminal proline residue is of utmost importance, and its mutation abolishes glycosylase activity completely, as it carries out the nucleophilic attack on the glycosidic bond of the damaged base (*28*). In addition, mutagenesis experiments have revealed that there are a number of Glu residues that affect Fpg glycosylase activity to various extents.

Notably, mutation of Glu-2 to Gln completely abolishes the glycosylase activity of Fpg (*50*). In fact, the Bst-Fpg E2Q mutant was utilized by Fromme and Verdine to obtain a crystal structure of a catalytically inactive Fpg with an intact 8-oxoG base within the active site region (*34*). In the wild-type *E. coli* enzyme in our simulations, the Glu-2 side chain carboxylic acid group resides in a region that essentially forms an oxyanion hole, accommodating the negative charge of the side chain with the main chain amide groups of Gly-167, Ile-169, and Tyr-170. In *anti* 8-oxodG simulation 1, the Oϵ2 group of Glu-2 participates in a hydrogen bond with Ile-169 for 99% of the simulation while the Oϵ1 group participates in hydrogen bonds with Gly-167 and Tyr-170 for 86 and 94% of the simulation, respectively. Through the introduction of an amino group and neutralization of the side chain by the E2Q mutation, these interactions are interrupted and Gln-2 of the Bst-Fpg enzyme hydrogen bonds to O4′ of the damaged base in the crystal structure (*34*). This shifts the sugar pucker to O4′-exo rather than O4′-endo, as observed in *anti* 8-oxodG simulation 1 (see Figure 2). This structural shift within the active site region could misalign the substrate base and side chain residues, resulting in the loss of glycosylase activity. The charge balance of the active site is skewed by the Glu → Gln mutation, which could prevent the substrate base moiety from leaving the sugar during catalysis, thus inhibiting cleavage of the glycosidic bond. In addition, reduction of the negative charge within the inner active site region, stemming from the Glu → Gln mutation, makes placing the electronegative O^8 group within this region more favorable, lending stability to the *syn* conformation of the damaged base (*35*). In the wild-type enzyme, the *anti* conformation may predominate since the inner active site pocket, near Glu-2, is more negatively charged, making the placement of the O^8 group in this region less favorable.

The E173Q mutation also completely abolishes the glycosylase activity of *E. coli* Fpg, but this residue is not in direct contact with the damaged base within the active site region. However, Glu-173 does directly hydrogen bond to Ile-169, which in turn hydrogen bonds to Glu-2; this is also absolutely necessary for glycosylase activity. It is likely that Glu-173 plays a crucial role in forming the active site pocket, albeit at a short distance, and mutation to Gln results in structural distortion of the residues directly involved in

recognition and catalysis, thus preventing glycosylase activity.

The E5Q mutation does not completely abolish glycosylase activity, but it does hamper, by ~3 times compared to that of the wild-type enzyme, the ability of Fpg to cleave the damaged base from the duplex (*50*). As discussed above, Glu-5 is involved in hydrogen bonding to the Watson−Crick edge of substrate bases that are excised, with the exception of FapyA, and its mutation to Gln may result in the partial disruption of this interaction. It is possible that either this activity is partially preserved after mutation to Gln or this function is not absolutely required for base cleavage. The E131Q mutation also reduces, but does not abolish, the glycosylase active site of Fpg (*50*). However, this residue is far-removed from the active site region and may play a role in the structural integrity of the protein or active site as a whole rather than a direct role in catalysis or base accommodation.

The Flexible Loop Region of Fpg Closes over the Substrate Base in the Active Site Region, and the Partner Base Is Stably Hydrogen Bonded to Arg-108. A number of X-ray crystal structures of Fpg enzymes from different prokaryotes are available, including structures of the apoprotein as well as liganded complexes (*29*, *34*, *51*−*53*). These structures have revealed the presence of a flexible loop region, which includes residues 217−224 in the *E. coli* enzyme. This loop was unresolved in the structure of *E. coli* Fpg covalently bound to DNA through a reduced Schiff base intermediate (*29*), which served as the starting structure for our modeling and simulations, as well as in *Lactococcus lactis* Fpg complexed with an abasic site-containing duplex. However, this loop was resolved in both the structure of the E2Q mutant Bst-Fpg, which contained the substrate 8-oxodG (*34*), and the structure of the uncomplexed *Thermus thermophilus* Fpg (*51*). Therefore, it appears that this flexible loop region is more mobile when the enzyme is complexed with DNA, but the substrate base has been excised, compared to the case when either the substrate base is present or the protein is not complexed with DNA. It is possible that the flexibility of this loop region increases upon cleavage of the glycosidic bond to allow the damaged base to exit the active site region and water to enter (*34*); it is thought that water participates in the β- and δ-elimination reaction steps that follow glycosidic bond cleavage (*53*).

The flexible loop, including residues 217−224, was modeled using homology modeling techniques (*29*, *35*), and the conformation of the loop following the simulations is similar to that seen in the crystal structure of the mutant Bst-Fpg complexed to 8-oxodG-containing duplex DNA (*34*). Figure 5 shows the overlaid loop regions of the Bst-Fpg crystal structure and our simulated structures; the alignments were carried out using the backbone atoms of protein residues 10−200 for each of the systems. Figure S4 shows the two-dimensional rmsd analysis of flexible loop region residues 217−224 in each simulated system. In each structure, Lys-217 resides closest to the damaged base and comprises one side of the active site pocket in which the substrate base resides. The homologous residue in Bst-Fpg is Arg-223, which also forms one side of the active site pocket and, as a basic residue, could act as a hydrogen bond donor. In the Bst-Fpg structure, the damaged base resides in the *syn* conformation, with the O^8 group out of reach of Arg-223. However, it is possible that the catalytic inactivity of the mutant enzyme results from a disruption of the appropriate interactions within the active site region, possibly shifting the *anti*−*syn* equilibrium. Perhaps the modified base may reside in either the *anti* or *syn* conformation prior to catalysis in a catalytically competent complex (*35*). If the 8-oxodG residue were to reside in the *anti* conformation within the active site of wild-type Bst-Fpg, it would likely be within hydrogen bonding distance of Arg-223, and Bst-Fpg could function in a manner analogous to that proposed for *E. coli* Fpg, with Arg-223 hydrogen bonding to the O^8 damage site of 8-oxodG.

In addition to interactions between the damaged base and the flexible loop region, interactions with the substrate base's partner have been shown to be important in excision by Fpg (*17*). Arg-108 has been deemed critical in discriminating against dA opposite the damaged base (*29*); the order of preference of the 8-oxodG partner is as follows: dT > dG > dC ≫ dA (*17*). During our simulations, Arg-108 was hydrogen bonded to the partner dC base in the unmodified, 8-oxodG, and FapydG simulations, as well as the dT base in the 8-oxodA and FapydA simulations. Figure S5 of the Supporting Information shows the distances and angles of these interactions during the simulations for each system. In the simulations with a dC partner, Arg-108 forms two or three hydrogen bonds with the partner base, including those between Nη2 of Arg-108 and N3 of dC, Nη2 of Arg-108 and O2 of dC, and Nϵ of Arg-108 and O2 of dC. When dT is located opposite the damaged base, as in the 8-oxodA and FapydA simulations, Arg-108 partakes in two hydrogen bonds with O2 of the partner base, through Nη2 and Nϵ.

Although dC partakes in two or three hydrogen bonds with the protein, as opposed to one or two between Fpg and dT, excision of 8-oxodG opposite dT is carried out more efficiently than that opposite dC. It is likely that dT and dG are preferred opposite the lesion because the 8-oxodG base is more easily flipped out of the helix into the active site region of Fpg. The damaged base may be more difficult to flip out of the helix when opposite dC because three hydrogen bonds are formed between the damaged base and its partner, adding to the energy barrier of base extrusion. In contrast, 8-oxodG does not form a stable base pair with either dT or dG in duplex DNA (*8*). It is likely that 8-oxodG is not excised when located opposite dA because the damaged base tends to stably adopt the *syn* conformation in duplex DNA when mismatched with dA (*8*); the adduct in the *syn* conformation would not present its O^8-damaged moiety in the major groove, a probable requirement for initial damage recognition and extrusion from duplex DNA into the active site region (*17*).

SUMMARY AND CONCLUSIONS

The mechanism by which distinct lesions are recognized and processed by Fpg has been of much interest in understanding how substrate specificity is accomplished. Namely, why are 8-oxoG, FapyG, and FapyA bound and excised by Fpg, while 8-oxoA and unmodified guanine are not excised? Specifically mysterious is the removal of FapyA while 8-oxoA opposite T is not excised; these bases differ only in the configuration of their imidazole rings. Our molecular modeling and molecular dynamics studies exam-

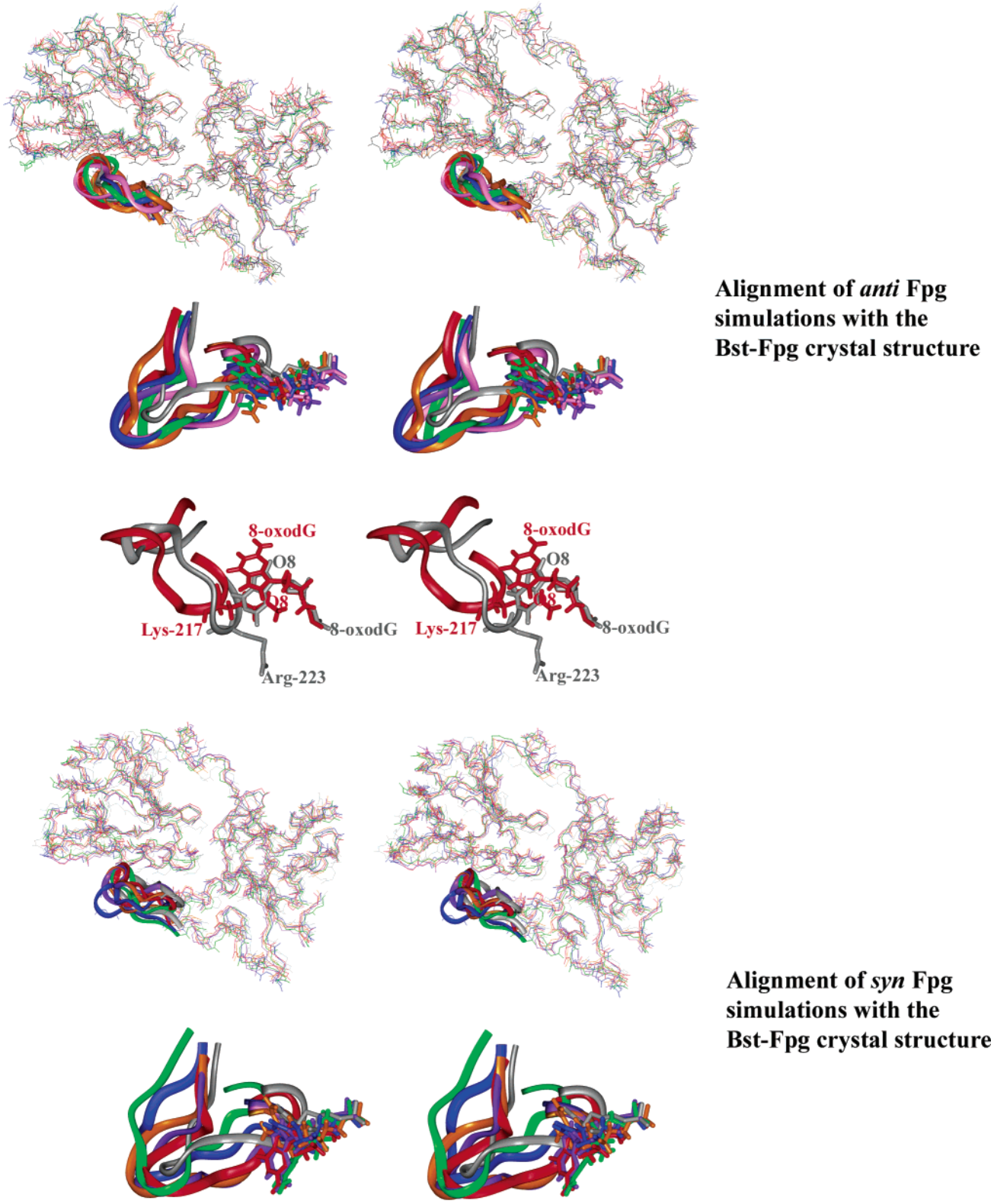

FIGURE 5: Alignment of the simulated systems (after 2 ns) with the crystal structure of Bst-Fpg (*34*), in which the flexible loop region was resolved. The color code follows: gray for the Bst-Fpg crystal structure, purple for the unmodified simulation, red for the 8-oxodG simulation, pink for *anti* 8-oxodG simulation 2, green for the FapydG simulation, orange for the FapydA simulation, and dark blue for the 8-oxodA simulation. Systems were aligned using the backbone atoms of residues 10−200. This figure includes the flexible loop regions (residues 215−228 for *E. coli* Fpg and residues 221−234 for Bst-Fpg), as well as the substrate base. In addition, the overlay of the Bst-Fpg crystal structure and *anti* 8-oxodG simulation 1 is shown, including Lys-217 (*E. coli*) and Arg-223, its homologous residue in the Bst-Fpg structure.

ined each of these five bases, unmodified guanine, 8-oxoG, FapyG, FapyA, and 8-oxoA, within the active site of *E. coli* Fpg in both the *anti* and *syn* conformations.

Both Lys-217 and Met-73 were pinpointed as candidates for damage recognition within the active site region; Lys-217 has been identified experimentally as being important

to substrate discrimination (*35*), while Met-73 could not be tested by mutagenesis experiments since its main chain carbonyl group is implicated. The 8-oxoG, FapyG, and FapyA bases interact with Lys-217 and/or Met-73 of Fpg at damage-specific positions in the *anti* orientation, and 8-oxoG also hydrogen bonds with Lys-217 in the *syn* orientation. However, neither FapyG nor FapyA interacts with either Lys-217 or Met-73 when the bases adopt the *syn* orientation, indicating that while 8-oxodG may be recognized in either the *anti* or *syn* conformation, the Fapy residues appear to be recognized by the enzyme only in the *anti* conformation. Interestingly, a crystal structure of an inactive carba-FapydG lesion bound to the Fpg protein from *L. lactis*, published while this paper was under review, also shows the carba-FapydG lesion adopting the *anti* conformation (*54*). This structure manifests many of the same interactions between the Fpg protein and FapyG moiety that we identify in the current simulation. Most notable is the water-mediated interaction between Met-73 (Met-75 in the *L. lactis* protein) and N7 and N9 of the damaged base.

Also of interest is a molecular dynamics investigation by Amara et al. (*55*), again published while this paper was under review; this work concluded that the *syn* conformation of 8-oxodG is favored within the Fpg active site based on the dynamic behavior of the flexible loop region. This computational result differs from the simulation reported by Zaika et al., which showed specific enzyme−DNA recognition for the *anti* conformation of 8-oxodG (*35*). While Amara et al. found that the *syn* conformation showed more water-mediated contacts with the loop, there were more direct contacts between the enzyme and lesion in their simulation of the *anti* conformation; furthermore, the loop was further from the 8-oxodG when it was *syn*. Thus, while Amara et al.'s hypothesis that loop mobility contributes to lesion recognition is an interesting concept, no evidence was presented to support their interpretation that the *syn* conformation is preferred simply because the loop is more flexible in this conformation. While loop dynamics might support the enzyme's lesion recognition function, it is equally, if not more, plausible that the stabilized loop region present in the simulation of *anti* 8-oxoG would more readily facilitate substrate stabilization and excision (*35*, *55*). Stop-flow kinetics of *E. coli* Fpg (*24*) shows that after 8-oxodG eversion and insertion of the void-filling residues, there is one more conformational change equilibrium step preceding base excision. In principle, this step could represent interconversion between *syn* and *anti* orientations of 8-oxodG. More likely, however, it corresponds either to a transition from a "prerecognition" conformation (in which the protein makes few contacts with the already-everted base, possibly reflecting loop disorder), to a "recognition" conformation (in which the protein makes the full set of contacts with the base), or to a transition from the recognition conformation to a "pre-excision" conformation in which some contacts are weakened, and base protonation (if any) occurs; this change also may be accompanied by loop disordering. It is possible that ordered and disordered loop structures could coexist in the reaction pathway while the base is still not excised.

The unmodified and 8-oxoA bases do not interact with Lys-217 or Met-73 in either the *anti* or *syn* conformations, consistent with Fpg's failure to cleave 8-oxoA from DNA. It is likely that although unmodified guanine is bound with a low affinity by Fpg due to its failure to present an O^8 group in the major groove of duplex DNA, the large abundance of guanine results in it being sampled within the active site of Fpg. The discriminating protein−DNA interactions depicted in these simulations elucidate how that secondary discrimination within the active site region occurs. Although 8-oxoA is likely extruded from the DNA duplex since it possesses an O^8 group, its glycosidic bond is not cleaved since it fails to fulfill the required interactions within the active site region, according to our simulated results (Figure 4). In contrast, while FapyA has a Watson−Crick edge identical to 8-oxoA, preventing some interactions with active site residues, like Glu-5 and Thr-214, the flexibility of the ring-opened imidazole ring allows the FapyA base moiety to hydrogen bond with Met-73, signaling to the protein that the substrate base is appropriate for excision.

These molecular dynamics simulations provide a foundation for future computational investigations of thermodynamics of binding to Fpg for these lesions. A pioneering experimental binding study for the 8-oxodG-Fpg system has already been carried out (*56*). In addition, the simulations present a basis for investigation of reaction pathways using hybrid molecular/quantum mechanical methods, as investigated in a related system (*57*). Finally, further computational and experimental investigations of mutant proteins will provide additional insights into these intriguing systems.

Elucidating the mechanism of substrate discrimination within the active site of Fpg yields a glimpse of the manner in which this prototype DNA glycosylase recognizes and excises a number of different substrates, while discriminating against structurally similar bases. This study also suggests that substrate bases may be recognized in more than one manner, namely, by hydrogen bonding to Lys-217 and/or Met-73, either directly or through water-mediated hydrogen bonds. Deciphering the mechanism by which glycosylase enzymes discriminate between appropriate and inappropriate substrate bases is important in understanding pathways that have evolved to protect cells from mutagenesis and other effects of DNA damage.

NOTE ADDED AFTER ASAP PUBLICATION

In the paragraph on this page that begins, "Also of interest", the sentence immediately following the second mention of ref *55* has been deleted. This paper was originally published 11/25/04; the corrected version was published 12/07/04.

SUPPORTING INFORMATION AVAILABLE

Five tables of atom types and partial charges, a table of parameters added to the Parm99.dat parameter set, and five figures depicting data for the simulations. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Halliwell, B., and Gutteridge, J. M. C. (1989) *Free Radicals in Biology and Medicine*, 2nd ed., Claredon Press, Oxford, U.K.
2. Marnett, L. J. (2000) Oxyradicals and DNA damage, *Carcinogenesis 21*, 361−370.
3. Shigenaga, M. K., Gimeno, C. J., and Ames, B. N. (1989) Urinary 8-hydroxy-2′-deoxyguanosine as a biological marker of in vivo oxidative DNA damage, *Proc. Natl. Acad. Sci. U.S.A. 86*, 9697−9701.

4. David, S. S., and Williams, S. D. (1998) Chemistry of glycosylases and endonucleases involved in base-excision repair, *Chem. Rev. 98*, 1221−1261.
5. Boiteux, S., Gajewski, E., Laval, J., and Dizdaroglu, M. (1992) Substrate specificity of the *Escherichia coli* Fpg protein (formamidopyrimidine-DNA glycosylase): excision of purine lesions in DNA produced by ionizing radiation or photosensitization, *Biochemistry 31*, 106−110.
6. Burrows, C. J., and Muller, J. G. (1998) Oxidative Nucleobase Modifications Leading to Strand Scission, *Chem. Rev. 98*, 1109−1152.
7. Breen, A. P., and Murphy, J. A. (1995) Reactions of oxyl radicals with DNA, *Free Radical Biol. Med. 18*, 1033−1077.
8. Shibutani, S., Takeshita, M., and Grollman, A. P. (1991) Insertion of specific bases during DNA synthesis past the oxidation-damaged base 8-oxodG, *Nature 349*, 431−434.
9. Grollman, A. P., and Moriya, M. (1993) Mutagenesis by 8-oxoguanine: an enemy within, *Trends Genet. 9*, 246−249.
10. Cheng, K. C., Cahill, D. S., Kasai, H., Nishimura, S., and Loeb, L. A. (1992) 8-Hydroxyguanine, an abundant form of oxidative DNA damage, causes G → T and A → C substitutions, *J. Biol. Chem. 267*, 166−172.
11. Pavlov, Y. I., Minnick, D. T., Izuta, S., and Kunkel, T. A. (1994) DNA replication fidelity with 8-oxodeoxyguanosine triphosphate, *Biochemistry 33*, 4695−4701.
12. Shibutani, S., Bodepudi, V., Johnson, F., and Grollman, A. P. (1993) Translesional synthesis on DNA templates containing 8-oxo-7,8-dihydrodeoxyadenosine, *Biochemistry 32*, 4615−4621.
13. Delaney, M. O., Wiederholt, C. J., and Greenberg, M. M. (2002) Fapy·dA induces nucleotide misincorporation translesionally by a DNA polymerase, *Angew. Chem., Int. Ed. Engl. 41*, 771−773.
14. Wiederholt, C. J., Greenberg, M. M., Haraguchi, K., Delaney, M. O., Sambandam, A., and Hantosi, Z. (2002) Fapy·dG instructs Klenow exo(−) to misincorporate deoxyadenosine, *J. Am. Chem. Soc. 124*, 7278−7279.
15. Scharer, O. D., and Jiricny, J. (2001) Recent progress in the biology, chemistry and structural biology of DNA glycosylases, *BioEssays 23*, 270−281.
16. Barnes, D. E., Lindahl, T., and Sedgwick, B. (1993) DNA repair, *Curr. Opin. Cell Biol. 5*, 424−433.
17. Tchou, J., Bodepudi, V., Shibutani, S., Antoshechkin, I., Miller, J., Grollman, A. P., and Johnson, F. (1994) Substrate specificity of Fpg protein. Recognition and cleavage of oxidatively damaged DNA, *J. Biol. Chem. 269*, 15318−15324.
18. Breimer, L. H. (1984) Enzymatic excision from γ-irradiated polydeoxyribonucleotides of adenine residues whose imidazole rings have been ruptured, *Nucleic Acids Res. 12*, 6359−6367.
19. Tchou, J., and Grollman, A. P. (1995) The catalytic mechanism of Fpg protein. Evidence for a Schiff base intermediate and amino terminus localization of the catalytic site, *J. Biol. Chem. 270*, 11671−11677.
20. Wiederholt, C. J., Delaney, M. O., and Greenberg, M. M. (2002) Interaction of DNA containing Fapy·dA or its C-nucleoside analogues with base excision repair enzymes. Implications for mutagenesis and enzyme inhibition, *Biochemistry 41*, 15838−15844.
21. Haraguchi, K., Delaney, M. O., Wiederholt, C. J., Sambandam, A., Hantosi, Z., Greenberg, M. M., and Rithner, C. D. (2002) Synthesis and characterization of oligodeoxynucleotides containing formamidopyrimidine lesions and nonhydrolyzable analogues: Studies on N4-(2-deoxy-D-pentofuranosyl)-4,6-diamino-5-formamidopyrimidine (Fapy·dA) and N6-(2-deoxy-D-pentofuranosyl)-6-diamino-5-formamido-4-hydroxypyrimidine (Fapy·dG), *J. Am. Chem. Soc. 124*, 3263−3269.
22. Plum, G. E., Grollman, A. P., Johnson, F., and Breslauer, K. J. (1995) Influence of the oxidatively damaged adduct 8-oxodeoxyguanosine on the conformation, energetics, and thermodynamic stability of a DNA duplex, *Biochemistry 34*, 16148−16160.
23. Ober, M., Linne, U., Gierlich, J., and Carell, T. (2003) The two main DNA lesions 8-oxo-7,8-dihydroguanine and 2,6-diamino-5-formamido-4-hydroxypyrimidine exhibit strongly different pairing properties, *Angew. Chem., Int. Ed. Engl. 42*, 4947−4951.
24. Koval, V. V., Kuznetsov, N. A., Zharkov, D. O., Ishchenko, A. A., Douglas, K. T., Nevinsky, G. A., and Fedorova, O. S. (2004) Pre-steady-state kinetics shows differences in processing of various DNA lesions by *Escherichia coli* formamidopyrimidine-DNA glycosylase, *Nucleic Acids Res. 32*, 926−935.
25. Wiederholt, C. J., Delaney, M. O., Pope, M. A., David, S. S., Greenberg, M. M., Haraguchi, K., Sambandam, A., Hantosi, Z., and Rithner, C. D. (2003) Repair of DNA containing Fapy·dG and its β-C-nucleoside analogue by formamidopyrimidine DNA glycosylase and MutY, *Biochemistry 42*, 9755−9760.
26. Nilsen, H., and Krokan, H. E. (2001) Base excision repair in a network of defence and tolerance, *Carcinogenesis 22*, 987−998.
27. Grollman, A. P., and Zharkov, D. O. (1999) Recognition and excision of bases from oxidatively damaged DNA by Fpg, Ogg1 and MutY proteins, in *Advances in DNA Damage and Repair* (Karakaya, D. A., Ed.) Plenum Publishers, New York.
28. Zharkov, D. O., Rieger, R. A., Iden, C. R., and Grollman, A. P. (1997) NH_2-terminal proline acts as a nucleophile in the glycosylase/AP-lyase reaction catalyzed by *Escherichia coli* formamidopyrimidine-DNA glycosylase (Fpg) protein, *J. Biol. Chem. 272*, 5335−5341.
29. Gilboa, R., Zharkov, D. O., Golan, G., Fernandes, A. S., Gerchman, S. E., Matz, E., Kycia, J. H., Grollman, A. P., and Shoham, G. (2002) Structure of formamidopyrimidine-DNA glycosylase covalently complexed to DNA, *J. Biol. Chem. 277*, 19811−19816.
30. Bruner, S. D., Norman, D. P., and Verdine, G. L. (2000) Structural basis for recognition and repair of the endogenous mutagen 8-oxoguanine in DNA, *Nature 403*, 859−866.
31. Lau, A. Y., Scharer, O. D., Samson, L., Verdine, G. L., and Ellenberger, T. (1998) Crystal structure of a human alkylbase-DNA repair enzyme complexed to DNA: mechanisms for nucleotide flipping and base excision, *Cell 95*, 249−258.
32. Slupphaug, G., Mol, C. D., Kavli, B., Arvai, A. S., Krokan, H. E., and Tainer, J. A. (1996) A nucleotide-flipping mechanism from the structure of human uracil-DNA glycosylase bound to DNA, *Nature 384*, 87−92.
33. Hollis, T., Ichikawa, Y., and Ellenberger, T. (2000) DNA bending and a flip-out mechanism for base excision by the helix-hairpin-helix DNA glycosylase, *Escherichia coli* AlkA, *EMBO J. 19*, 758−766.
34. Fromme, J. C., and Verdine, G. L. (2003) DNA lesion recognition by the bacterial repair enzyme MutM, *J. Biol. Chem. 278*, 51543−51548.
35. Zaika, E. I., Perlow, R. A., Matz, E., Broyde, S., Gilboa, R., Grollman, A. P., and Zharkov, D. O. (2004) Substrate discrimination by formamidopyrimidine-DNA glycosylase: A mutational analysis, *J. Biol. Chem. 279*, 4849−4861.
36. Pearlman, D. A., Case, D. A., Caldwell, J. W., Ross, W. S., Cheatham, T. E. I., DeBolt, S., Ferguson, D. M., Seibel, G. L., and Kollman, P. A. (1995) AMBER, a package of computer programs applying molecular mechanics, normal mode analysis, molecular dynamics, and free energy calculations to simulate the structural and energetic properties of molecules, *Comput. Phys. Commun. 91*, 1−41.
37. Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K., Ferguson, D., Spellmeyer, D., Fox, T., Caldwell, J., and Kollman, P. A. (1995) A second generation force field for the simulation of proteins, nucleic acids, and organic molecules, *J. Am. Chem. Soc. 117*, 5179−5197.
38. Cheatham, T., Cieplak, P., and Kollman, P. (1999) A modified version of the Cornell et al. force field with improved sugar pucker phases and helical repeat, *J. Biomol. Struct. Dyn. 16*, 845−862.
39. Darden, T. A., York, D., and Pedersen, L. (1993) Particle Mesh Ewald: An N·log(N) method for Ewald sums in large systems, *J. Chem. Phys. 98*, 10089−10092.
40. Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., and Bourne, P. E. (2000) The Protein Data Bank, *Nucleic Acids Res. 28*, 235−242.
41. Mao, H., Deng, Z., Wang, F., Harris, T. M., and Stone, M. P. (1998) An intercalated and thermally stable FAPY adduct of aflatoxin B1 in a DNA duplex: structural refinement from ^{1}H NMR, *Biochemistry 37*, 4374−4387.
42. Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., Zakrzewski, V. G., Montgomery, J. A., Stratmann, R. E., Burant, J. C., Dapprich, S., Millam, J. M., Daniels, A. D., Kudin, K. N., Strain, M. C., Farkas, O., Tomasi, J., Barone, V., Cossi, M., Cammi, R., Mennucci, B., Pomelli, C., Adamo, C., Clifford, S., Ochterski, J., Petersson, G. A., Ayala, P. Y., Cui, Q., Morokuma, K., Salvador, P., Dannenberg, J. J., Malick, D. K., Rabuck, A. D., Raghavachari, K., Foresman, J. B., Cioslowski, J., Ortiz, J. V., Baboul, A. G., Stefanov, B. B., Liu, G., Liashenko, A., Piskorz, P., Komaromi, I., Gomperts, R., Martin, R. L., Fox, D. J., Keith, T., Al-Laham, M. A., Peng, C. Y., Nanayakkara, A., Challacombe, M., Gill, P. M. W., Johnson, B., Chen, W., Wong, M. W., Andres, J. L.,

Gonzalez, C., Head-Gordon, M., Replogle, E. S., and Pople, J. A. (2001) *Gaussian 98*, revision A.10, Gaussian, Inc., Pittsburgh, PA.

43. Bayly, C., Cieplak, P., Cornell, W., and Kollman, P. (1993) A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model, *J. Phys. Chem. 97*, 10269−10280.
44. Wu, X., Shapiro, R., and Broyde, S. (1999) Conformational analysis of the major DNA adduct derived from the food mutagen 2-amino-3-methylimidazo[4,5-*f*]quinoline, *Chem. Res. Toxicol. 12*, 895−905.
45. Hingerty, B., Ritchie, R., Ferrell, T., and Turner, J. (1985) Dielectric effects in biopolymers: the theory of ionic saturation revisited, *Biopolymers 24*, 427−439.
46. Mezei, M. (1997) Simulaid: A collection of utilities designed to help setting up molecular dynamics simulations, *J. Comput. Chem. 18*, 812−815.
47. Reddy, C. K., Das, A., and Jayaram, B. (2001) Do water molecules mediate protein-DNA recognition? *J. Mol. Biol. 314*, 619−632.
48. Tchou, J., Kasai, H., Shibutani, S., Chung, M. H., Laval, J., Grollman, A. P., and Nishimura, S. (1991) 8-Oxoguanine (8-hydroxyguanine) DNA glycosylase and its substrate specificity, *Proc. Natl. Acad. Sci. U.S.A. 88*, 4690−4694.
49. Boiteux, S., O'Connor, T. R., Lederer, F., Gouyette, A., and Laval, J. (1990) Homogeneous *Escherichia coli* FPG protein. A DNA glycosylase which excises imidazole ring-opened purines and nicks DNA at apurinic/apyrimidinic sites, *J. Biol. Chem. 265*, 3916−3922.
50. Lavrukhin, O. V., and Lloyd, R. S. (2000) Involvement of phylogenetically conserved acidic amino acid residues in catalysis by an oxidative DNA damage enzyme formamidopyrimidine glycosylase, *Biochemistry 39*, 15266−15271.
51. Sugahara, M., Mikawa, T., Kumasaka, T., Yamamoto, M., Kato, R., Fukuyama, K., Inoue, Y., and Kuramitsu, S. (2000) Crystal structure of a repair enzyme of oxidatively damaged DNA, MutM (Fpg), from an extreme thermophile, *Thermus thermophilus* HB8, *EMBO J. 19*, 3857−3869.
52. Serre, L., Pereira de Jesus, K., Boiteux, S., Zelwer, C., and Castaing, B. (2002) Crystal structure of the *Lactococcus lactis* formamidopyrimidine-DNA glycosylase bound to an abasic site analogue-containing DNA, *EMBO J. 21*, 2854−2865.
53. Fromme, J. C., and Verdine, G. L. (2002) Structural insights into lesion recognition and repair by the bacterial 8-oxoguanine DNA glycosylase MutM, *Nat. Struct. Biol. 9*, 544−552.
54. Coste, F., Ober, M., Carell, T., Boiteux, S., Zelwer, C., and Castaing, B. (2004) Structural basis for the recognition of the FapydG lesion (2,6-diamino-4-hydroxy-5-formamidopyrimidine) by the Fpg DNA glycosylase, *J. Biol. Chem*. (in press).
55. Amara, P., Serre, L., Castaing, B., and Thomas, A. (2004) Insights into the DNA repair process by the formamidopyrimidine-DNA glycosylase investigated by molecular dynamics, *Protein Sci. 13*, 2009−2021.
56. Minetti, C. A., Remeta, D. P., Zharkov, D. O., Plum, G. E., Johnson, F., Grollman, A. P., and Breslauer, K. J. (2003) Energetics of lesion recognition by a DNA repair protein: thermodynamic characterization of formamidopyrimidine-glycosylase (Fpg) interactions with damaged DNA duplexes, *J. Mol. Biol. 328*, 1047−1060.
57. Dinner, A. R., Blackburn, G. M., and Karplus, M. (2001) Uracil-DNA glycosylase acts by substrate autocatalysis, *Nature 413*, 752−755.

BI048747F